

The Structure of Goal Systems Predicts Human Performance

David D. Bourgin^{1a} (ddbourgin@berkeley.edu)

Falk Lieder^{1a} (falk.lieder@berkeley.edu)

Daniel Reichman^{1a} (daniel.reichman@gmail.com)

Nimrod Talmon^{1b} (nimrodtalmon77@gmail.com)

Thomas L. Griffiths^a (tom_griffiths@berkeley.edu)

^a Department of Psychology
University of California Berkeley
Berkeley, CA 94720 USA

^b Department of Computer Science and Applied Mathematics
Weizmann Institute of Science
Rehovot, Israel

¹ These authors contributed equally.

Abstract

Most psychological theories attribute people's failure to achieve their goals exclusively to insufficient motivation or lack of skill. Here, we offer a complementary explanation that emphasizes the inherent complexity of the computational problems that arise from the structure of people's goal systems. Concretely, we hypothesize that people's capacity to achieve their goals can be predicted from combinatorial parameters of the structure of the network connecting their goals to the means available to pursue them. To test this hypothesis, we expressed the relationship between goals and means as a bipartite graph where edges between means and goals indicate which means can be used to achieve which goals. This allowed us to map two computational challenges that arise in goal achievement onto two classic NP-hard problems: Set Cover and Maximum Coverage. The connection between goal pursuit and NP-hard problems led us to predict that people should perform better with goal systems that are tree-like. Three behavioral experiments confirmed this prediction. Our results imply that network parameters that are instrumental to algorithm design could also be useful for understanding when and why people struggle in their goal pursuits.

Keywords: decision-making; goals; rational analysis; graph theory; computational complexity

Introduction

The ability to set and achieve high-level goals, such as creating a CogSci paper, is a critical feature of human intelligence and a key challenge for artificial intelligence systems (Newell & Simon, 1972). Critically, everyday problem solving requires people to juggle multiple goals in parallel (Atkinson & Birch, 1970; Miller, Galanter, & Pribram, 1960). Concretely, when people are given ten minutes to list their current pursuits they will report about 15 goals on average and each of those goals typically entails multiple subgoals at several levels of abstraction (Little & Gee, 2007).

It is generally agreed that there are many situations in which people fail to act on their goals (Baumeister, Heatherston, & Tice, 1994). The predominant explanations of such failures are lack of motivation, lack of planning (Gollwitzer, 1999), failure to delay gratification (Mischel, Shoda, & Rodriguez, 1989), or the depletion of the capacity for self-control (Muraven & Slessareva, 2003). Here, we explore an alternative explanation for people's failure to achieve their

goals: the inherent complexity of the underlying computational problem.

The relationship between means and goals can be formalized in a bipartite graph whose vertices are divided into a set of means $M = \{m_1, \dots, m_k\}$ and a set of goals $G = \{g_1, \dots, g_l\}$. In this graph, a mean m is connected to a goal g if and only if selecting m will achieve the goal g . Such networks are called *goal systems* in the psychological literature (Kruglanski et al., 2002). For instance, the vertices at the top of the goal system illustrated in Figure 1 might correspond to your goals to become a prolific scientist (g_1), be a wonderful partner (g_2), become a great parent (g_3), get physically fit (g_4), and enjoy life to the fullest (g_5).

Finding the best configuration of means for achieving a set of goals can involve considerable computational challenges. It has been suggested that findings from theoretical computer science can shed light on how people cope with hard computational problems (Van Rooij, 2008). For example, Van Rooij (2008) advocated applying the theory of fixed parameter tractability to study how people cope with hard computational problems. Yet, while previous research on problem solving has investigated which strategies people use to solve NP-hard problems (MacGregor & Ormerod, 1996; MacGregor, Ormerod, & Chronicle, 2000), this literature has focused on the Traveling Salesman problem and other problems that are structurally distinct from those that arise in goal pursuit.¹ Here, we will fill this gap by analyzing goal achievement through the lens of computational complexity theory.

In theoretical computer science, it is well known that the performance of many combinatorial optimization algorithms critically depends on certain graph-theoretic properties of the networks they are applied to (Kleinberg & Tardos, 2006). For instance, a well-documented phenomenon in algorithm design, artificial intelligence, and operational research is that NP-hard optimization problems often become easier on trees and tree-like graphs. Indeed, when restricted to trees, many

¹One exception is the work of Carruthers, Masson, and Stege (2012) which found that the planarity of graphs has no effect on human performance in the Vertex Cover problem.

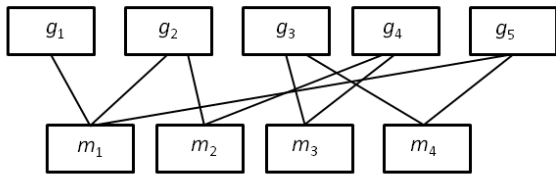


Figure 1: Example of goal system: means appear at the bottom marked by $m_1 \dots m_4$. Goals appear on top marked by $g_1 \dots g_5$. A goal can be attained if at least one mean connected to it is chosen. The set $\{m_1, m_3\}$ is a minimal set of means covering all goals

NP-hard problems can be solved by efficient polynomial-time algorithms, such as divide-and-conquer methods and greedy algorithms. Here, we show that human performance on two goal management tasks is also well predicted by graph theoretic measures of the tree-likeness of the underlying goal system. Our results offer a fresh computational perspective on why people fail to achieve their goals. Our experimental results align well with theoretical knowledge from computer science and highlight that findings from computational complexity are relevant to cognitive psychology.

The plan for the paper is as follows. We start by formalizing two common challenges of goal achievement in terms of two classic NP-hard problems: Set Cover and Maximum Coverage. Next we derive our theoretical prediction by arguing that people’s performance on these problems should increase with graph theoretic metrics of how tree-like the goal system is. We then test this prediction in three behavioral experiments and conclude with a summary of our findings and directions for future work.

Formal analysis and predictions

Here we formalize, as well-defined NP-hard problems, two computational challenges that arise in means selection problems where one seeks to choose a set of means that are instrumental to the ends one is trying to achieve.

The first problem we consider is trying to achieve as many goals as possible with a fixed budget that limits the number of activities that one can perform. We formalize this challenge in terms of the Maximum Coverage problem (MC). In it, we are given a bipartite graph $H = (C, D, F)$ where C, D are the sides of the bipartition, and F is the set of edges connecting vertices in C to vertices in D . We are also given a nonnegative integer $k \leq |C|$. We seek to find a set $C' \subseteq C$, of cardinality at most k , maximizing the number of vertices in D covered by vertices in C' (a vertex $b \in D$ is covered by a vertex $a \in C$ if $(a, b) \in F$). As a goal system, the set C corresponds to means, the set D corresponds to goals, and F represents to interconnections between goal and means. Observe that we assume that once a goal is covered by a *single* mean then it will be achieved. This assumption is made in order to simplify the experimental task, allowing for a simple and clean

description.

The second problem that we study is trying to achieve a given set of goals as efficiently as possible by selecting the minimal number of means that will accomplish all goals. We formalize this challenge in terms of the Set Cover problem (SC). In the Set Cover problem, we are given a bipartite graph $G = (A, B, E)$, where A, B are the sides of the bipartition and E is the set of edges connecting vertices in A to vertices in B . As a goal system, the set A corresponds to means, the set B corresponds to goals, and the set E corresponds to interconnections between goals and means. In the SC problem, our goal is to cover all vertices in B by a set $A' \subseteq A$ of minimal cardinality. Both MC and SC are NP-hard not only to solve exactly but also to approximate (Feige, 1998).

These two computational problems (MC and SC) capture essential aspects of means selection problems that have been studied in the psychological literature.

For example, it is generally agreed that people try to select means in order to maximize the number of attained goals (Kruglanski et al., 2002; Zhang, Fishbach, & Kruglanski, 2007). Furthermore, representing goals as graphs and assuming interconnectedness between goals appear either explicitly or implicitly in several papers (Kruglanski et al., 2002; Thagard & Millgram, 1995).

We shall use the following two performance measures in quantifying how well people do in MC and SC. Our first measure simply quantifies the ability to find the optimal solution. This is a binary measure that equals 1 if the person finds the optimal solution and 0 otherwise. Given an algorithm A for MC or SC, the second measure, referred to as the *solution quality*, ranges between 0 (worst) and 1 (best). For the MC task, the solution quality is the number of goals achieved by A divided by maximum number of goals that could be achieved with k means. For the SC task, the solution quality is the minimum number of means that achieves all goals divided by the number of means selected by A . Solution quality (which is also referred to as the *approximation ratio* of A) is widely used in quantifying the quality of approximation algorithms for NP-hard problems (Vazirani, 2013).

The hardness of MC and SC gives a first indication for why people might find it difficult to juggle multiple goals at the same time. Yet, not every instance of these problems is equally difficult. Next, we introduce graph-theoretic measures that might be useful for distinguishing harder instances from easier ones.

Features and predictions

A tree is a connected graph without cycles. Many NP-hard problems on graphs with small treewidth (Robertson & Seymour, 1986, see below), allow exact or approximate algorithms which are significantly better than what is known to be achievable on worst-case instances.

The idea that tree-like graphs might be easier for people to deal with guided our search for features that quantify how similar a given network is to a tree. Four such features are presented below.

- *Treewidth.* Treewidth (tw) is a combinatorial parameter that is associated with a graph.² Low treewidth implies that the nodes and edges of the network can be arranged in a way that resembles a tree (e.g., Kloks, 1994; Robertson & Seymour, 1986). For a precise definition of treewidth see Kloks (1994).
- *Combinatorial expansion.* Given a graph $G = (V, E)$ and a nonempty subset $S \subseteq V$, let $\partial(S)$ be the set of edges crossing the cut $(S, V \setminus S)$ and let $N(S)$ be the set of all vertices in $V \setminus S$ having a neighbor in S . The vertex-expansion of G is defined as $\min_{S \subseteq V, 0 < |S| \leq |V|/2} \frac{|N(S)|}{|S|}$. The edge-expansion of G is $\min_{S \subseteq V, 0 < |S| \leq |V|/2} \frac{|\partial(S)|}{|S|}$. Trees of bounded degree have expansion $O(1/|V|)$, hence large expansion suggests that the graph is dissimilar to a tree. We computed both vertex and edge expansion by solving an integer linear programs (ILPs) using IBM’s CPLEX.
- *Spectral expansion.* The adjacency matrices of the graphs we consider are symmetric, hence have n real eigenvalues $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$. The classical (discrete) Cheeger’s inequality (e.g., Alon and Milman, 1985) implies that the larger $d - \lambda_2$ is, the larger the edge expansion of the graph.³

MC and SC instances with low treewidth w are known to have exact algorithms that run in time $O(2^w n)$ (Alber & Niedermeier, 2002), hence instances with low treewidth are likely to be easier to deal with algorithmically. Current algorithms that compute tree-decompositions are quite complicated, therefore it seems unlikely that people would use them to solve SC and MC problems. Nevertheless, treewidth might affect the performance of people’s heuristics: the similarity between low treewidth graphs and trees might make the kinds of algorithms that people might use, such as greedy and divide-and-conquer methods, much more effective. Conversely, as worst-case instances of MC and SC are hard even to approximate, and as hard instances often have large treewidth and expansion (Clementi & Trevisan, 1999), it is likely that it will be hard not only to solve exactly, but also to find approximate solutions for instances of large treewidth. Similar reasoning applies to our expansion measures. We therefore hypothesize that treewidth, vertex-expansion, edge-expansion, and the spectral gap of G are negatively correlated with the quality of people’s solution to SC and MC problems and frequency with which they find an optimal solution.

Additional predictors

A feedback vertex set (FVS) is a subset of vertices whose removal from a given graph results in a forest. The size of a minimal feedback vertex set is an alternative measure to the similarity of a graph to a tree. Hence we used this

²Here we will only consider undirected graphs.

³We included the spectral gap as it admits an efficient polynomial algorithm as opposed to the other three predictors which are NP-hard. As such, it may prove as a practical predictor of performance.

feature as well. We calculated FVS using the implementation based on (Iwata, 2016; Wahlström, 2014) available at <https://github.com/wata-orz/fvs>.

Previous empirical hardness models have found additional features of graphs to be useful: the diameter, average eccentricity, and average path length (Leyton-Brown, Nudelman, & Shoham, 2009). We thus included these features. The diameter of a graph is the longest distance between two vertices in the graph (where the distance, $dist(u, v)$, is the number of edges in a shortest path connecting u and v ; all graphs considered are connected). The eccentricity $\epsilon(v)$ of a vertex v in an undirected graph $G = (V, E)$ is the maximal distance of a vertex in $V \setminus v$ from v . The average eccentricity (AvgEcc) is $\frac{1}{n} \sum_{v \in V} \epsilon(v)$ (n is the number of vertices). The average path length (AvgPath) is $\frac{2}{n(n-1)} \sum dist(u, v)$ where summation is taken over all pairs of distinct vertices.

Behavioral experiments

To test how our predictors relate to human performance in MC and SC, we conducted three crowdsourced behavioral experiments. We used a between-subjects design where each participant was assigned randomly to one of twenty graphs with treewidths varying from 4 to 13. In the first experiment, participants were asked to solve the SC problem, and in the second experiment participants were asked to solve the MC problem. In each case, the problem was graphically represented as a bipartite graph with 48 vertices. The 24 vertices at the bottom represent the available means (activities A-Z) and 24 vertices at the top represent the goals. Each edge from a means vertex to a goal vertex implies that completing that activity is sufficient to achieve the goal. In the SC task, participants were asked to select a minimal number of activities to achieve all of the goals. In the MC task, participants were asked to choose five activities that achieve as many goals as possible. The third and final experiment asked participants to solve a SC problem where goals are given semantic content and real values, and a different visual display is used to eliminate possible visualization effects. In each experiment we restricted our analyses to goal systems in which every goal and mean vertex had exactly 4 neighbors (i.e., each graph was 4-regular). This restriction meant that each graph required the same amount of memory to enable processing, ensuring that any difference between conditions cannot be explained by working memory limitations.

Experiment 1: Human performance on Set Cover

Methods We recruited 655 participants on Amazon Mechanical Turk. Participants were paid \$1.25 and could earn a performance-dependent bonus of up to \$2. Each participant was randomly assigned to one of 20 conditions that differed only in the graph structure of the SC problem participants were asked to solve. After consenting to participate, participants read a cover story about a person trying to choose which set of activities (e.g., volunteer to improve the company’s website and work out at the gym) they should perform in order to achieve all their goals (e.g., earn more money, improve

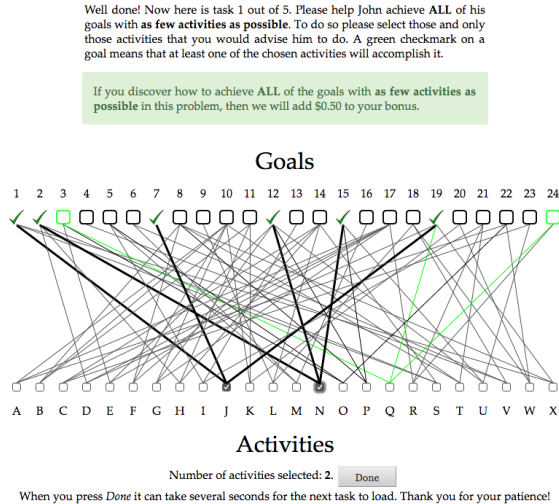


Figure 2: Interface and instructions for Experiment 1.

relationship with boss, get fit, etc.) with as few activities as possible because their time is limited. The story highlighted that some activities achieve multiple goals at the same time. Next, participants completed a simple practice trial involving only two means and two goals. Once they had solved this task successfully, participants proceeded to the Set-Cover problem they had been assigned to. When participants moused over an activity, this interface highlighted the goals it would achieve and the corresponding edges of the graph in green. Goals that had already been achieved and activities that had already been selected were highlighted by checkmarks, and the number of selected activities was shown at the bottom of the screen. Participants were asked to help a person achieve their 24 goals with as few activities as possible, and they were motivated by the prospect of earning a financial bonus of \$2 for achieving all goals as efficiently as possible. After submitting their solution, participants completed an exit questionnaire asking them for basic demographic information (age, gender, and primary language), and four 9-point Likert scales (anchors: “not at all”, “somewhat”, and “extremely”) measuring the perceived difficulty of the task, motivation to achieve all goals, motivation to find the optimal solution, and their motivation to finish the task as quickly as possible. We excluded 30 participants (4.6%) because their responses did not achieve all goals suggesting that they did not follow the instructions.

Results We found that treewidth alone explained 44.90% of the variance in the frequency with which people found the optimal solution across the 20 graphs ($F(1, 18) = 14.20, p = 0.0012$): as we increased the treewidth of the graph the percentage of participants who discovered the optimal solution decreased significantly ($\rho = -0.59, p = 0.0058$) from more than 90% on the graph with treewidth 5 to only about 30% on the graph with treewidth 14. We found that the aver-

age solution quality was negatively correlated with treewidth ($\rho = -0.44, p = 0.0525$) suggesting that our participants achieved fewer goals for goal systems with higher treewidth. Treewidth explained 17.59% of the variance in the median response time across problems ($F(1, 18) = 3.86, p = 0.0650$): the median amount of time people took to solve the problems tended to increase with treewidth ($\rho = 0.3426$) but this effect was not statistically significant ($p = 0.1393$), and when we restricted this analysis to correct solutions the correlation was $\rho = 0.3825$ ($p = 0.1297$). Perceived difficulty also tended to increase with treewidth ($\rho = 0.37$) but this correlation was not statistically significant ($p = 0.1062$). Our participants were highly motivated to find the optimal solution (average rating 7.91 ± 0.06 out of 9). Thus it appears unlikely that their motivation was a bottleneck to their performance. Furthermore, motivation appeared to be unaffected by treewidth ($\rho = -0.23, p = 0.34$). Thus the observed differences in performance appear to result from the inherent difficulties of the means selection problems posed by different goal systems.

Of the additional predictors evaluated, we found that graph diameter, average shortest path, and average graph eccentricity all were significantly positively correlated with the frequency of optimal solutions identified by our participants (graph diameter: $\rho = 0.5691, p = 0.0088$, avg. shortest path: $\rho = 0.6516, p = 0.0019$, avg. eccentricity: $\rho = 0.6265, p = 0.0031$). In addition, the spectral expansion (measured as $d - \lambda_2$) and the size of the graph vertex and edge expansions showed significant negative correlations with the frequency of optimal solutions (vertex expansion: $\rho = -0.5836, p = 0.0069$, edge expansion: $\rho = -0.4552, p = 0.0437$, spectral expansion: $\rho = -0.6280, p = 0.0030$). We also found that the average shortest path, average graph eccentricity, spectral expansion, and the size of the graph vertex expansions were significantly correlated with the average participant solution qualities (avg. shortest path: $\rho = 0.4505, p = 0.0462$, avg. eccentricity: $\rho = 0.4316, p = 0.0574$, spectral expansion: $\rho = -0.4496, p = 0.0467$, vertex expansion: $\rho = -0.4636, p = 0.0395$). Finally, only the cardinality of the graph edge expansions exhibited a significant correlation with the median response times on the SC task ($\rho = 0.4538, p = 0.0445$).

Experiment 2: Human performance on Maximum Coverage

Methods Experiment 2 was identical to Experiment 1 except for the task: participants were now instructed to achieve as many goals as possible subject to the constraint that the person’s limited time does not permit them to complete more than five activities. The 20 graphs and financial incentives were the same as in Experiment 1. The interface of Experiment 1 was modified to prevent participants from selecting more than five activities at a time. When a participant attempted to add a sixth activity they were told they would first have to remove one or more of the activities they had already selected. The cover story and survey were modified slightly to match the change in the task. We recruited 545 participants on Amazon Mechanical Turk. Participants were paid \$1.25

and could earn a bonus of up to \$2. The consent form specified that participants must not have participated in the previous version of this experiment. We excluded 23 participants (4.2%) because they had selected fewer than five means.

Results The frequency with which people found the optimal solution decreased significantly with treewidth ($\rho = -0.4828, p = 0.0311$). We found that treewidth alone explained 20.41% of the variance in the frequency with which people found the optimal solution across the 20 graphs ($F(1, 18) = 4.62, p = 0.0455$). We found that treewidth explained 25.25% of the variance in solution quality ($F(1, 18) = 6.08, p = 0.0240$) which significantly deteriorated as treewidth increased ($\rho = -0.4972, p = 0.0257$). The median amount of time people took to solve the problems did not increase significantly with treewidth ($\rho = 0.25, p = 0.28$) and treewidth explained only 0.6% of our participants' median response times ($F(1, 18) = 0.10, p = 0.76$). When we restricted the analysis to the time taken by optimal solutions, the relationship was still not statistically significant ($\rho = 0.2994, p = 0.1998; F(1, 18) = 0.10, p = 0.76$). Finally, treewidth explained only 8.8% of the variance in the perceived problem difficulty across the 20 graphs ($F(1, 18) = 1.74, p = 0.20$), and the correlation between treewidth and perceived difficulty was not statistically significant ($\rho = 0.26, p = 0.26$). Our participants were highly motivated to find the optimal solution (average rating 8.11 ± 0.06 out of 9). Thus it appears unlikely that their motivation was a bottleneck to their performance. Furthermore, motivation appeared to be unaffected by treewidth ($\rho = -0.03, p = 0.91$). Thus the observed differences in performance appear to result from the inherent difficulties of the means selection problems posed by different goal systems.

In addition, we found that both the size of the graph edge expansion and the graph spectral expansion (measured again as $d - \lambda_2$) were significantly negatively correlated with the frequency of optimal solutions (edge expansion: $\rho = -0.4802, p = 0.0321$, spectral expansion: $\rho = -0.4782, p = 0.0330$), while the average shortest path and average graph eccentricity showed a significant positive relationship (avg shortest path: $\rho = 0.4912, p = 0.0279$, avg. eccentricity: $\rho = 0.4391, p = 0.0528$). In contrast, only the size of the graph vertex and edge expansions showed a significant correlation with the average solution quality (vertex expansion: $\rho = -0.4431, p = 0.0504$, edge expansion: $\rho = -0.4832, p = 0.0309$), suggesting that in general graph treewidth and combinatorial expansions may be more robust predictors of human performance on the MC problem. None of the metrics surveyed were significantly correlated with median participant response times.

Experiment 3: A more realistic Set-Cover task

While Experiments 1 and 2 capture some of the computational challenges of goal achievement, the tasks were relatively abstract. Experiment 3 addresses this limitation by as-

signing semantic labels to the 24 goals. These labels were common new-years resolutions such as “get in shape” and “earn more money”. Similar semantic goals were used in previous research in goal-system theory (Zhang et al., 2007). We also used a different interface to avoid possible visualization effects that arise from graph drawings in the first two experiments.

Methods We recruited 600 participants on Amazon Mechanical Turk. Participants were paid \$0.38 for about 5min of work plus a performance-dependent bonus of \$0.50 if they found an optimal solution. Each participant was randomly assigned to one of the twenty graph structures used in Experiments 1 and 2. For each graph, the order in which the means were listed and the order in which the goals were listed was randomized between participants. The participants' task was to achieve all goals with as few means as possible. The graphical interface of the task was changed to reduce visual clutter. Instead of drawing edges between mean and goals, the goals achieved by each mean were listed next to it (see Figure 3). The cover story was similar to the one used in Experiment 1 but the training trial used the new task interface shown in Figure 3. The consent form required that participants had not participated in any of our previous goal management experiments. All participants were included in the subsequent analyses.

Results On a scale from 1 to 9 participants rated their motivation to find a solution that achieves all goals with the minimal number of means as 7.38, their motivation to finish the task as quickly as possible and move on as 4.35, and the difficulty of the task as 5.67. We found that treewidth, the magnitude of the graph spectral expansion, cardinalities of the graph edge and vertex expansions, average eccentricity, average shortest path, and graph diameter were all significantly correlated with the frequency with which human participants identified the optimal solution (treewidth: $\rho = -0.756, p = 0.0001$; avg. eccentricity: $\rho = 0.583, p = 0.007$; avg. shortest path: $\rho = 0.651, p = 0.002$; graph diameter: $\rho = 0.525, p = 0.017$; spectral expansion: $\rho = -0.708, p = 0.0005$; edge expansion: $\rho = -0.7, p = 0.0006$; vertex expansion: $\rho = -0.7303, p = 0.0003$). Similarly, treewidth, the magnitude of the graph spectral expansion, cardinalities of the graph edge and vertex expansions, average eccentricity, and average shortest path were all significantly correlated with the average solution quality of human responses (treewidth: $\rho = -0.60, p = 0.005$; avg. eccentricity: $\rho = 0.4872, p = 0.0293$; avg. shortest path: $\rho = 0.5243, p = 0.0176$; spectral expansion: $\rho = -0.5308, p = 0.0160$; edge expansion: $\rho = -0.5670, p = 0.0091$; vertex expansion: $\rho = -0.6047, p = 0.0047$). None of the features were significantly correlated with median participant response times.

Well done! Now here is the real task: Please help John achieve ALL of his goals with as few activities as possible. To do so please select those and only those activities that you would advise him to do. A green checkmark on a goal means that at least one of the chosen activities will accomplish it.

If you discover how to achieve ALL of the goals with as few activities as possible, then we will pay you a bonus of \$0.5.

Activities	Serves Goals	Goal
<input type="checkbox"/> A	12, 18, 21, 23	<input type="checkbox"/> 1. Volunteer/give to charity
<input type="checkbox"/> B	2, 3, 9, 24	<input type="checkbox"/> 2. Improve mental skills (e.g., concentration)
<input type="checkbox"/> C	3, 8, 12, 21	<input type="checkbox"/> 3. Start a hobby
<input type="checkbox"/> D	8, 9, 15, 24	<input type="checkbox"/> 4. Learn about art, music, and culture
<input type="checkbox"/> E	3, 5, 12, 18	<input type="checkbox"/> 5. Get a better job
<input type="checkbox"/> F	1, 6, 9, 15	<input type="checkbox"/> 6. Get in shape
<input type="checkbox"/> G	11, 17, 20, 22	<input type="checkbox"/> 7. Start saving money

Figure 3: Graphical Interface of Experiment 3.

Conclusions

We demonstrated that people's performance in Maximum Coverage and Set Cover can be reliably predicted from graph theoretic measures for the tree-likeness of the goal system, such as treewidth and expansion. Our data support the conclusion that tree-like goal systems are easier for people to handle. More generally, our results imply that parameters that are used in theoretical computer science to differentiate between hard and easy instances can be leveraged to predict human performance in NP-hard tasks.

One limitation of our experiments is that their complete, explicit representation of goals, means, and the connections between them is a simplifying idealization. In real life, people are often unaware of some of their goals and means, as well as some of the connections between goals and means. For example, maintaining goal systems of moderate size in working memory when solving means selection problems is likely to be nontrivial. Hence real-life representations of goals are likely to make means selection problems as those discussed here even more challenging to solve. Although such memory problems are not directly related to how tree-like the goal system is, they are nevertheless consistent with our hypothesis that the cognitive difficulty of means selection is an important limiting factor for people's ability to achieve their goals.

In conclusion, our results suggest that even highly motivated people will likely fall short of achieving all their goals when they have to consider many goals and means in parallel. Our analyses provide a novel approach to predicting how likely people are to succeed in these settings, with implications for the design of goal systems that make it easier for people to meet their objectives.

References

Alber, J., & Niedermeier, R. (2002). Improved tree decomposition based algorithms for domination-like problems. In *Latin american symposium on theoretical informatics* (pp. 613–627).

Atkinson, J. W., & Birch, D. (1970). *The dynamics of action*. Oxford: John Wiley.

Baumeister, R. F., Heatherton, T. F., & Tice, D. M. (1994). *Losing control: How and why people fail at self-regulation*. San Diego: Academic Press.

Carruthers, S., Masson, M. E., & Stege, U. (2012). Human performance on hard non-euclidean graph problems: vertex cover. *The Journal of Problem Solving*, 5(1), 5.

Clementi, A. E., & Trevisan, L. (1999). Improved non-approximability results for minimum vertex cover with density constraints. *Theoretical Computer Science*, 225(1), 113–128.

Feige, U. (1998). A threshold of $\ln n$ for approximating set cover. *Journal of the ACM*, 45(4), 634–652.

Gollwitzer, P. M. (1999). Implementation intentions: strong effects of simple plans. *American Psychologist*, 54(7), 493–503.

Iwata, Y. (2016). Linear-time kernelization for feedback vertex set. *arXiv preprint arXiv:1608.01463*.

Kleinberg, J., & Tardos, É. (2006). *Algorithm design*. Pearson Education India.

Kloks, T. (1994). *Treewidth: computations and approximations* (Vol. 842). Springer Science & Business Media.

Kruglanski, A. W., Shah, J. Y., Fishbach, A., Friedman, R., Chun, W. Y., & Sleeth-Keppler, D. (2002). A theory of goal systems. *Advances in Experimental Social Psychology*, 34, 331–378.

Leyton-Brown, K., Nudelman, E., & Shoham, Y. (2009). Empirical hardness models: Methodology and a case study on combinatorial auctions. *Journal of the ACM*, 56(4), 22.

Little, B. R., & Gee, T. (2007). The methodology of personal projects analysis: Four modules and a funnel. In B. R. Little, K. Salmela-Aro, & S. D. Philipps (Eds.), *Personal project pursuit: Goals, action, and human flourishing*. New York: Psychology Press.

MacGregor, J. N., & Ormerod, T. (1996). Human performance on the traveling salesman problem. *Perception & Psychophysics*, 58(4), 527–539.

MacGregor, J. N., Ormerod, T. C., & Chronicle, E. (2000). A model of human performance on the traveling salesperson problem. *Memory & Cognition*, 28(7), 1183–1190.

Miller, G. A., Galanter, E., & Pribram, K. H. (1960). *Plans and the structure of behavior*. New York: Adams Bannister Cox.

Mischel, W., Shoda, Y., & Rodriguez, M. L. (1989). Delay of gratification in children. *Science*, 244(4907), 933–938.

Muraven, M., & Slessareva, E. (2003). Mechanisms of self-control failure: Motivation and limited resources. *Personality and Social Psychology Bulletin*, 29(7), 894–906.

Newell, A., & Simon, H. A. (1972). *Human problem solving*. Englewood Cliffs: Prentice-Hall.

Robertson, N., & Seymour, P. D. (1986). Graph minors. ii. algorithmic aspects of tree-width. *Journal of algorithms*, 7(3), 309–322.

Thagard, P., & Millgram, E. (1995). Inference to the best plan: A coherence theory of decision. In A. Ram & D. B. Leake (Eds.), *Goal driven learning* (pp. 439–454). MIT Press.

Van Rooij, I. (2008). The tractable cognition thesis. *Cognitive Science*, 32(6), 939–984.

Vazirani, V. V. (2013). *Approximation algorithms*. Springer Science & Business Media.

Wahlström, M. (2014). Half-integrality, LP-branching and FPT algorithms. In *Proceedings of the twenty-fifth annual acm-siam symposium on discrete algorithms* (pp. 1762–1781).

Zhang, Y., Fishbach, A., & Kruglanski, A. W. (2007). The dilution model: how additional goals undermine the perceived instrumentality of a shared path. *Journal of Personality and Social Psychology*, 92(3), 389–401.